CHAPTER II. PROBLEMS AND METHODS

Most of the problems of constructing production indexes arise from the fact that the available data are incomplete and are inadequate in certain other respects. The particular problems encountered and the methods used to resolve them are discussed in the following sections.

Lack of Data on Materials Consumption

As mentioned earlier, gross output measures for individual industries were used as estimates of the desired net measures because it was not found possible to calculate the latter directly. The reason for this is the unavailability, in the census or elsewhere, of comprehensive data on quantities of the various materials used in manufacture, the "Q's" of formula (3), and on the values of these materials from which might be derived the "P's." Quantity, and in many cases value, figures are available in the census for certain selected materials consumed in 1947, 1939, or both years, but this information is too limited to be of general use in calculating net output indexes.

There are two possible causes of difference between indexes of gross and net output at constant prices for an individual industry. The first relates to changes in the proportion of total work done on a product of an industry that is done in the industry itself. For instance, materials may be purchased which are more or less highly fabricated than previously—or are of better or poorer quality—thus saving operations that would otherwise have been performed in the industry, or requiring previously unnecessary operations. Such changes in effect represent a shift of part of the total work done from or to the given industry, and would not be reflected in the gross measure, which relates only to the change in the number of units of product turned out. More efficient use of materials may result in reducing the amount required per unit of product, with a consequent higher net output for the same amount of gross output. Technological change may make possible the substitution of less expensive materials without affecting the quality of the product, again leading to more net output per unit of gross. Changes of these types, which are more likely to be important the longer the time span under consideration, will result in differences between the calculated gross measures and the desired net measures.

The second possible cause of difference between gross and net measures for an industry arises if the industry makes several products, and the ratio of value to value added per unit of product [1] differs for the various products. In such cases, a change in the "product mix"—the relative amounts of the different products made—may lead to differences in the movements of net and gross output. For example, an increase in the output of one product might be exactly offset in terms of the change in *value* at constant prices by an equivalent decrease in the output of another product whose price (in the constant price system) is the same. But if the value added per unit of the first product is higher than that of the second, the two changes would not offset in terms of the *value added* measure; net output would have risen.

This second type of error (but not the first type) can be avoided by using value added per unit, rather than price, as the multiplier

[1] Strictly speaking, "value added" cannot be calculated for individual products, but only for establishments and, by addition, for industries and combinations thereof. The term "value added per unit of product" implies an allocation of joint materials costs, such as those for common raw materials, fuels, etc., among the various products of the establishment, and such allocations often can be made only in an arbitrary fashion. The concept is useful, nevertheless, and where joint inputs are of small importance, value added per unit can be calculated (given the necessary data) with only a minor element of arbitrariness.

of quantities in formula (4). It was found possible in two industries (meat-packing and primary nonferrous metals) to make estimates of value added by type of product, and to use these rather than unit values in constructing the industry indexes. For the most part, however, this could not be done, and both potential sources of error exist in the individual industry indexes as estimates of net output change.

The importance of the errors introduced in a particular index by using gross industry measures as estimates of net cannot be determined without the very data which makes estimation necessary. Errors resulting from changes in the product mix can as readily be in one direction as the other, and undoubtedly are in both directions. There is some hope that they will offset each other—at least partially—when combinations into aggregate indexes are made. At best this is a hope, however, for there is no guarantee of offset in any particular aggregate. The errors resulting from change over time in the ratio of net to gross output for a particular product made in an industry can also be in both directions; but to the extent that they are the result of increased efficiency in materials use or of other sorts of technological change, understatements of the net output rise are likely to predominate. It is probable, therefore, that most of the aggregate indexes are biased downward by a presently undeterminable amount, although the error in any particular group or industry index may be negligible or in the other direction.[2]

Quality Changes

Many of the individual products treated as identical in the two years compared are really not identical, because of quality changes. Continuing research, invention, and discovery over the years have led to a more or less steady improvement in the quality of most manufactured products. These changes are difficult if not impossible to measure quantitatively, and the failure to reflect them in the present index, as well as in previously calculated indexes, results in a downward bias.

Another kind of "quality change" can result from the fact that groups of commodities treated statistically as individual products frequently change in composition. All women's dresses, for instance, are grouped together in the present calculations, in spite of the fact that dresses of many different prices and qualities are produced. If the proportion of more expensive dresses was higher in 1947 than in 1939, the average quality of the category "dresses" may be said to have increased. Such changes in the proportions of various price lines do not always occur, of course, and with regard to many products they may have been in the other direction. To avoid the potential errors on this score, as far as possible the most detailed breakdowns of product types

[2] The question of possible differences between net and gross output measures for Canadian industry is currently being investigated at the Dominion Bureau of Statistics, under the direction of Mr. V. R. Berlinguette. Canadian industrial censuses, which are taken annually, differ from those of the United States in that rather comprehensive data are collected for many industries on quantities and costs of materials consumed in manufacture. The Bureau of Statistics is constructing both net and gross annual output measures from census data for as many industries as possible, for the period since 1935. This work is still in process, but preliminary results for selected industries, mainly manufactured foods, indicate that (1) differences between the net and gross measures do occur, and while often they are comparatively small, occasionally they are quite large; (2) the differences are in both directions, but—at least in the industries so far investigated—the net output measures, on balance, tend to be higher than the gross measures; and (3) changes in product mix are a more important cause of difference for the industries and time periods studied than changes in the ratios of net to gross output for particular products. These conclusions, of course, are based on partial results and are consequently highly tentative. The final results should throw considerable light on this general subject.

that the basic data permitted were used. In a few cases it was found possible to estimate breakdowns of 1939 figures, from closely related figures, to match the greater detail available for 1947. Despite these efforts to achieve homogeneity, it is probable that the unavoidable combinations of related commodities, like true quality changes in individual commodities, result in some bias.

Lack of Quantity Data for Many Commodities

Probably the most important single difficulty encountered in making production indexes is that they must be compiled without physical volume data for a considerable number of products. There are several reasons for this. For some types of manufacturing output (e. g., aeronautical development work) there exists no acceptable physical unit of measure. For others (e. g., hardware), the differences among individual commodities produced are so great that a count of the number of items has little meaning. In some instances (e. g., furniture) a substantial number of producers are unable to report physical volume data with reasonable accuracy. Certain "miscellaneous" and "not reported by kind" items included in the census are, of course, not reported in quantity terms. Efforts to improve commodity classifications frequently result in destroying comparability with figures for earlier periods. New products usually are not reported in physical volume terms until they have attained a position of economic importance, at which time there are no corresponding figures for earlier periods.

Every effort was made to supplement the product data of the Census of Manufactures by use of statistics from other sources, such as the Bureau of Mines, the Department of Agriculture, the Tariff Commission, and the Bureau of Internal Revenue. (In a few of the cases where data were available from both census and noncensus sources, the noncensus data were used because they appeared superior.) About 1,700 products, listed in appendix A, were separately included in the present indexes.[3] This compares with 837 products utilized in Fabricant's indexes for 1937 (relative to the year 1929). The difference represents some increase in the coverage of available quantity data, but is mainly one of the degree of detail; in the present calculation the products included covered 60 percent of the total value of all manufactured products, and in Fabricant's they covered 51 percent.[4]

The problems posed by the lack of physical quantity data for the remaining products can best be considered in two parts. One part concerns the industries for which the available quantity data were so limited that they were considered an insufficient basis for the development of reasonably reliable industry indexes. The treatment of these is described in the later section, "Adjustments for missing industries." The other concerns the industries for which the available data, while usually not complete, were considered adequate for making first approximations to industry indexes—the "unadjusted" indexes. To these first approximations the so-called "industry coverage adjustment" was applied as a means of taking account of the remaining products not represented by quantity data. This adjustment, which was developed by Mills and used by Fabricant, is intended also to correct simultaneously for another problem—the fact that the available product data for each industry usually include some production originating outside of the industry in question. Because many establishments make products "primary" to more than one industry, statistics for the total output of a particular product type cannot usually be identified exactly with the output of the industry primarily making this type of product. The census quantity figures show total output of the product types "primary" to each

[3] This number was obtained simply by counting the "product lines" on the index work sheets. In many cases these represent combinations of related products as defined in the 1947 census, so that the actual number of "products" included is substantially higher.

[4] In both indexes, certain products were represented by data for materials consumed in their manufacture. See "Alternative Methods Used for Some Industries." If these are included the percentages of total value covered are 66 and 55, respectively.

industry, including the output of these products made in other industries.

The industry coverage adjustment is discussed in detail in appendix D in connection with both of these problems. Briefly, the adjustment rests on the assumption that the average change in price between the 2 years for all of the industry's production was the same as the average price change that can be calculated for the part of the production, which also includes some output in other industries, for which quantity data are available. If the adjustment is not made, and the unadjusted indexes are taken to represent "industry" output, an alternative assumption is implicit: that the average quantity change of all products made in the industry was the same as the average quantity change calculated from incomplete quantity data, which partially overlap output in other industries. The assumption underlying the adjusted indexes is believed to be superior, partly on the basis of certain tests.

The adjusted indexes are calculated in effect by deflating the value of products made in the industry by a price index developed for represented products primary to the industry. Actually, a different but mathematically equivalent method was employed which involved first calculating unadjusted indexes, and then correcting them by a "coverage adjustment factor" which is the quotient of the "coverage ratios" for the 2 years. The specific definitions of these terms, the rationale of the procedure, and the tests that were made of its validity are discussed in appendix D.

Both the adjusted and unadjusted industry indexes are shown in chapter IV, together with the coverage ratios and adjustment factors. In addition to indicating the magnitude of the adjustment, the unadjusted indexes are useful as measures of output change for particular categories of products without regard to industry of origin.

The "Standard" Method Used for Industry Indexes

The general procedures used for calculating the industry indexes can be summarized in terms of a "standard" method. Under this method the data used were taken primarily from tables 1 and 6 of the industry reports of the 1947 Census of Manufactures, Volume II, *Statistics by Industry*. In table 1 of each report are given general industry statistics for recent census years, including, among other things, information on value of products, value added, and number of employees, summarized for all establishments of the industry. These statistics are republished in appendix B of this report, including revisions, special adjustments, and estimates made after publication of the 1947 census basic volumes, mainly in the course of constructing the present indexes.

Table 6 of the industry reports shows quantity and value of individual products made (or shipped) in 1947 and 1939. To obtain the unadjusted index for an industry, those products were selected from table 6 for which comparable quantity and value figures were available for 1947 and 1939, and the data posted to a work sheet, in as fine detail as the requirement of comparability permitted. These data were supplemented by the following, also posted to the work sheet:

(1) Other product data of table 6 which are not comparable as published, with estimated adjustments to make them comparable, involving such things as estimates of quantities corresponding to the part of the output of a product for which only values were reported, deduction of excise taxes from 1939 product values where these are included, adjustments where some producers reported product values at the distributor or retail level rather than f. o. b. factory, conversion of data on a fiscal or crop year basis to the calendar year, allowance for known undercoverage in either year, and so forth;

(2) Unpublished census data, some of which were not separately shown in the census volumes to avoid disclosing the operations of individual companies;

(3) Data published in the 1939 Census of Manufactures volumes, but not republished in the 1947 volumes;

(4) Data from monthly and quarterly census surveys; and

(5) Data published by other Government agencies.

The quantity and value data used are shown in appendix A, insofar as they can be published. All adjustments and supplements to census data as originally published in the 1947 basic volumes are described in appendix C.

For the industries for which the Bureau of Labor Statistics also calculates indexes,[5] the quantity figures used in the census indexes generally are those agreed upon as best in a joint review of the available statistics made by the staffs of the two agencies. The BLS and census industry indexes are not always identical, however, because of differences in methods of combining the basic data that follow from the differences in objectives described earlier.

Unit values for each product for both years were derived by dividing values of products by the respective quantities. Three unadjusted indexes were then calculated for each industry, using 1939, 1947, and the average of 1939 and 1947 unit values as weights, according to the following formulas in which, in the symbolism of appendix D, the primes indicate that the data relate to represented primary products wherever made:

(1) Unadjusted industry index, 1939 weights: $\dfrac{\Sigma q'_{47} p'_{39}}{\Sigma q'_{39} p'_{39}}$

(2) Unadjusted industry index, 1947 weights: $\dfrac{\Sigma q'_{47} p'_{47}}{\Sigma q'_{39} p'_{47}}$

(3) Unadjusted industry index, cross weights: $\dfrac{\Sigma q'_{47}(p'_{39}+p'_{47})}{\Sigma q'_{39}(p'_{39}+p'_{47})}$

To calculate the coverage ratios, the values of the individual products posted to the work sheet were added for 1947 and also for 1939, and each of the sums was divided by the total value of products made in the industry in the corresponding year, as shown in table 1 of the industry reports. The adjustment factor was obtained by dividing the 1947 coverage ratio by the 1939 coverage ratio. (These terms are explained in appendix D.) Each of the three unadjusted indexes was then divided by this adjustment factor to obtain the adjusted indexes.

Preliminary unadjusted and adjusted indexes were calculated in this manner for all industries. The indexes were retained in the final calculations if the coverage ratio in both years was at least 40 percent.[6] Where the coverage ratio fell below 40 percent, the available output data were considered to be inadequate for obtaining reasonably reliable indexes, and the preliminary indexes were discarded.[7] In the majority of retained indexes the coverage ratios were considerably in excess of 40 percent, as is indicated in the frequency distribution of these ratios shown in table 4.

Special Problems in Connection with the Industry Coverage Adjustment

The coverage adjustment for individual industries requires comparable industry value of products figures for the 2 years. The data shown in table 1 of the industry reports are usually although not always comparable. The difficulties encountered and the procedures used to overcome them are described below. As will be noted, some of these difficulties also involve the product data used for calculating the unadjusted indexes.

Shipments versus production.—In table 6 of the various industry reports of the 1947 census, volume II, the columns for 1947 usually are headed "total shipments and interplant transfers—quantity, value" and the columns for 1939, "total production for sale and interplant transfer—quantity, value." Similarly, table 1 of the reports shows the "value of products shipped" by the industry for 1947, and the "value of products made" for 1939. These headings correspond to the types of information called for

on the reporting forms in the 2 years, and if the figures reported were as requested, the data for the 2 years would be noncomparable to the extent that production and shipments differed. It is believed, however, that the data for the most part are directly comparable, because many manufacturers probably reported shipments in 1939 even though production was requested, and because the changes in finished goods inventories in most industries were of minor importance and there was, therefore, little difference between production and shipments in either 1939 or 1947. Under the standard method the published figures were treated as comparable.

TABLE 4.—FREQUENCY DISTRIBUTION OF INDUSTRY INDEXES, BY COVERAGE RATIO, FOR 1947 AND 1939

Coverage ratio	Number of indexes	
	1947	1939
Total	217	217
Below 40.0	8	6
40.0–49.9	10	5
50.0–59.9	12	11
60.0–69.9	19	16
70.0–79.9	28	25
80.0–89.9	35	43
90.0–99.9	64	67
100.0–109.9	19	20
110.0–119.9	5	1
120.0 and over	7	9
Coverage ratio not computed [1]	10	14

[1] Coverage ratios were not computed where the data necessary to make a coverage adjustment were not available, and in cases where the index was based on data for materials consumed. Coverage ratios were computed for 1947 in the cases of four industries for which 1939 ratios could not be computed, although no coverage adjustments were made for these industries.

NOTE.—A frequency distribution of coverage adjustment factors (calculated, for each industry, by dividing the 1947 coverage ratio by that for 1939) is shown in table D-2 of appendix D. Both the ratios and the factors are shown for individual industries in ch. IV.

In those industries in which, prior to the census canvass, it seemed likely that shipments would differ significantly from production in 1947, data on quantities produced were also requested on the reporting forms and are shown in a supplementary column of table 6. In making the indexes for these industries the production rather than shipments data for individual products in 1947 were used (with unit values derived from quantities and values of products shipped) except in certain cases where (1) the difference was less than 2 percent, or (2) the difference was mainly due to the fact that a part of the reported production was used for further manufacture in the producing plant and was not produced for sale or transfer.

For most of the industries in which establishments were asked to report data on production as well as shipments of individual products, special calculations of the industry's value of products in 1947 were made in compiling the census statistics.[8] Where the calculated value of products for the industry differed from the value of shipments by more than 2 percent, the former was used in making the coverage adjustment.

Duplication.—In a number of industries the products of some establishments are sold or transferred to other establishments of the same industry. In compiling the 1947 census, the values of such intra-industry sales were generally excluded from the value of shipments data if the products were resold by the receiving establishments without further processing. This was not done in the 1939 census, however, so that the 1939 value of products data for the industries in which such "merchandising" occurred are overstated relative to the statistics for 1947. It will be recalled that under the standard method the adjusted industry indexes are calculated (in effect) by dividing the change in industry value

---

[5] See ch. I, "Definition of Physical Volume."

[6] The requirement of a 40 percent coverage ratio also was used in the National Bureau study. For an analysis of its implications, see Fabricant, *The Output of Manufacturing Industries, 1899–1937*, pp. 364–66.

[7] With a few exceptions, mainly in the stone, clay, and glass, and in the machinery (except electrical) industry groups where indexes were retained for three and four industries, respectively, for which the coverage ratios in one or both years were 25 to 40 percent. These indexes were used in calculating the group indexes, but, with one exception, are not shown separately in the tables.

[8] These figures were not published, but were used to derive the published figures on value added (by subtracting cost of materials from value of products) for the affected industries. See the 1947 census basic volumes, "General Explanations," paragraph headed "Method of Calculating Value Added by Manufacture for Selected Industries in 1947."

of products by a price index, in order to obtain a measure of the quantity change. To the extent that the 1939 value of products data are overstated on account of merchandising, the adjusted indexes for 1947 relative to 1939 would be too low. Estimates were therefore made of the amounts of such resales in 1939 for the industries in which they occurred, and these were subtracted from the published 1939 values of products before calculating the adjusted industry indexes.

Intra-industry sales also occur in a number of industries for the purpose of further processing of the products sold, and the amounts of such sales were not excluded from the industry value data of either census. (Value of shipments figures were not published in the 1947 census for 12 industries in which the proportion duplicated was known to exceed 10 percent.) Since this sort of duplication affects the data for both years, it creates no problem so long as the proportion that duplicated shipments are of the total is the same in the 2 years, for in such cases the percentage change in shipments including duplication would agree exactly with the percentage change net of duplication. However, it is possible in individual cases for this relationship to have changed between 1939 and 1947 as a result of changes in the organization of production. For example, a given company in 1939 may have fabricated parts and assembled them into final products in the same establishment. The company's report for that year would relate only to the final products, for the parts were not made for sale as such. By 1947, however, the company may have begun to fabricate parts in a separate establishment (classified in the same industry) or, perhaps, to purchase them from other establishments in the industry, and the values of the parts would then be included in industry shipments. The 1947 shipments figures would, as a result, be overstated (or understated, if the opposite had happened) relative to 1939, and the adjusted indexes would be incorrect.

Because of the possibility of changes having occurred in the proportion of duplication in the industries where it was substantial, estimates of its amount in both years were subtracted from industry shipments [9] wherever it was found possible to do so. This was done, for example, in the steel works and rolling mills industry with regard to semifinished shapes and forms, and in the fertilizer industry with regard to superphosphates sold for making mixed fertilizers. Other cases are described in the industry notes of appendix C. Where no basis could be found for estimating the amount of duplication, a decision was made as to whether it was likely that its importance had changed substantially between 1939 and 1947. Where this appeared probable (e. g., motor vehicles) no adjusted indexes were calculated; where it did not (e. g., refrigeration equipment, radios and related products) the value of shipments data were used as reported in calculating the adjusted indexes.

A special case of duplication arises in a number of apparel and a few textile and leather industries, in connection with work done on contract by one establishment for another in the same industry. Both the value of the contract work done and the value of the final product (including the value of work done on contract) are included in the industry's value of shipments. Where the problem was important, the amount of duplication was estimated from census data on receipts and payments for contract work in each industry, and subtracted from the value of shipments in both years.

Establishment "splits."—For census purposes an establishment is defined as a single plant or factory, and is not necessarily identical with the business unit or company which may consist of one or more establishemnts. There has been some variation in the application of this definition in recent censuses, particularly with regard to companies carrying on distinctly different lines of activity at a single location. Specifically, in the 1947 census canvass there was some tendency to "split" 1939 establishments

(covered by a single report in that year) into two or more establishments corresponding to the different lines of activity and to obtain separate reports for each. This was done primarily in the tobacco products and paper products groups, and in some sections of the metals and metal products (including machinery) industries.

Establishment "splits" of this sort affect the comparability of the 1939 and 1947 value of products figures, and consequently the accuracy of the adjusted indexes. For example, in the tobacco products group, some companies primarily engaged in making cigarettes also make substantial quantities of chewing and smoking tobacco at the same location. In the 1939 census, consolidated reports were received for all activities at these locations, and the total value of products made there, including the value of chewing and smoking tobacco produced (a "secondary" product), was assigned to the cigarette industry. In the 1947 census, separate reports were obtained for cigarette production and chewing and smoking tobacco production at these locations, and these reports were classified in the respective industries. The consequence is that the increase in value of products from 1939 to 1947 shown in the census for the "cigarette" industry is too small, since the 1939 figure includes much chewing and smoking tobacco production not included in 1947, and the increase shown for the chewing and smoking tobacco industry is too large.

There are two possible means for overcoming this noncomparability in the value of products data. One is to combine the industries affected by the split, where the figures are comparable in total although not by parts. This was done in the case of tobacco. The other method is to construct 1939 value of products figures on a basis comparable with those for 1947 for the two industries affected by each split. This can be done by dividing the product values shown on the affected companies' original 1939 reports among the industries in the manner in which they would have been classified had separate reports been obtained in 1939. This method is feasible only where the necessary product detail is shown in the 1939 report, and where the number of plants involved is relatively small. It was followed in connection with the splits of malleable iron foundries from automobile plants, steam turbine generator set manufacture from motors and generators plants, and in one or two other cases.

Where neither of these methods was considered feasible, a decision was made as to whether the split significantly affected the comparability of the value of products statistics. If the data were substantially comparable in spite of the split, the coverage adjustment was made. If the amount of noncomparability was considered serious, only an unadjusted index was calculated for the industry.

Other Problems in Individual Industries

Classification changes.—The Standard Industrial Classification System was used in compiling the 1947 census with minor modifications listed in appendix B to the basic census volumes. This classification system differs substantially from that originally used for the 1939 census. For the most part this did not produce major problems in constructing the indexes because the product data, being on a "wherever made" basis, are largely unaffected by the industry classification system used, and the industry data (value of products, value added, etc.) for 1939 were retabulated according to the 1947 classification system in preparing the 1947 census volumes.

In some industries, however, the change in classification did produce problems which required special treatment. For example, the tobacco stemming and redrying industry was considered to be a manufacturing industry and was included in the census in 1947 but not in 1939. Efforts to estimate 1939 statistics for this industry were unsuccessful, and it was necessary to omit it from the index calculations. The dairy products industries were redefined in the 1947 census to exclude dairy products made by fluid-milk distributors. Data were collected and published on both the old and new bases for 1947, but were available only on the old basis for 1939. For index purposes it was necessary,

---

[9] Products known to be shipped to other establishments of the industry for further fabrication or resale were also omitted from the product detail used in calculating the unadjusted indexes, where data on the amounts were available or could be estimated.

therefore, to use the old rather than the new definitions of these industries. The concrete products industry was redefined to exclude ready-mixed concrete, but in retabulating the 1939 industry statistics it was not found possible to exclude all activity in connection with this product. Estimated adjustments to the published 1939 data were therefore made for the index. The treatment of similar problems in other industries is described in appendix C.

Combinations of industries.—As noted earlier, only combined indexes were calculated for some industries as a means of meeting the problem of establishment splits (e. g., cigarettes and chewing and smoking tobacco). Combinations were also made where the degree of product overlap was very great (e. g., the meat-packing and prepared-meats industries), and where alternative methods of representing output were used, as discussed below (e. g., all industries in the printing and publishing group). Finally, certain other combinations were made for special reasons, given in appendix C.

Alternative Methods Used for Some Industries

Indexes were calculated from data on quantities of output for 241 industries covering 68 percent of the value added in 1947. For a number of additional industries, data were available or could be estimated for both years on the quantities of a major raw material consumed. Where there was no evidence that important shifts had occurred in the types of materials employed, the change in materials consumption was taken as a measure of the change in output.[10] This was done for the entire printing and publishing group, and for five other individual industries, mainly fabricators of metals. For one industry, truck trailers, value data were deflated by a price index as a means of estimating the physical output change. In all, some 22 additional industries were represented by these means, raising the proportion of 1947 value added covered by specific indexes from 68 to 75 percent. This information and comparable data for Fabricant's 1929–37 indexes are summarized in table 5.

TABLE 5.—NUMBER AND IMPORTANCE OF REPRESENTED INDUSTRIES, BY METHOD OF REPRESENTATION, FOR PRESENT 1939–47 INDEXES AND FOR FABRICANT'S 1929–37 INDEXES

Method of representation	Number of indexes		Number of Industries		Percent of total value added	
	1939–47	1929–37	1939–47	1929–37	1939–47	1929–37
Grand total			453	(¹)	100.0	100.0
Total represented by specific indexes	217	143	263	(¹)	75.4	60.4
Product data	210	142	241	(¹)	67.6	53.4
Materials consumed	6	1	21	(¹)	7.7	7.0
Deflated value of products	1		1	(¹)	0.1	0.0
Not represented by specific indexes			190	(¹)	24.6	39.6

¹ Not shown because industry definitions used in Fabricant's indexes and in the 1947 Census of Manufactures were substantially different.

NOTE.—"Number of industries" exceeds "Number of indexes" because in some cases a single index was used to represent two or more industries. See "Combinations of industries."

Calculation of Group Indexes

Preliminary group indexes (i. e., unadjusted for "missing" industries) were calculated from the available industry indexes in each group, according to the following formulas, in which "I_{39}" and "I_{47}" (expressed as ratios rather than as percentages) represent the adjusted industry indexes for 1947 (relative to 1939)

with 1939 and 1947 weights, respectively, and "VA" the value added in the subscript period:

(1) Unadjusted group index, 1939 weights: $\dfrac{\Sigma(I_{39} \cdot VA_{39})}{\Sigma VA_{39}}$

(2) Unadjusted group index, 1947 weights: $\dfrac{\Sigma VA_{47}}{\Sigma(VA_{47} \div I_{47})}$

(3) Unadjusted group index, cross weights [11]: $\dfrac{\Sigma(VA_{47} + VA_{39} \cdot I_{39})}{\Sigma(VA_{39} + VA_{47} \div I_{47})}$

The value added figures used for combining industry indexes are given in appendix B. As noted there, the figures shown are identical with those published in the 1947 census basic volumes, except where modifications were made to improve the comparability of the 1939 and 1947 data. Most of these modifications in the value added figures were counterparts of revisions in the value of shipments data which were discussed earlier in connection with the problems of establishment "splits," classification changes, etc.

Since value added is calculated as the difference between cost of materials and value of shipments, any errors in these figures, or any noncomparabilities in the data for the 2 years, will also affect the value added statistics. In industries characterized by integrated operations, and in which, therefore, "interplant transfers" are important in value of shipments, cost of materials, or both, there may be some degree of noncomparability in the 1939 and 1947 value added data because of differences in the 2 years in the manner of reporting these operations, and because of the general difficulties associated with assigning transfer values.[12] Similarly, some manufacturers may have incorrectly included the cost of supplies for plant construction and repair in "cost of materials," which would result in understatements in value added. It is likely that such understatements, if they exist to a significant extent, would occur in the industries in which construction work was important. If these happened also to be the industries in which the increase in product output was relatively great, the indexes would have a downward bias on this account, as a result of underweighting the faster growing industries. It was not found possible to evaluate possible errors of these sorts in the value added data, but it is believed that they generally are not of great importance.

Adjustments for missing industries.—As is indicated in table 5, the missing industries in total represented about 25 percent of the value added by manufacture in 1947. Their number and importance are shown by major industry group in table 6, and they are individually identified in appendix B by the code letter "N."[13]

A number of the missing industries are the so-called "not elsewhere classified" industries, usually consisting of establishments making products of insufficient importance individually to call for separate industry classifications—the "left-overs" after the more important products in the subgroup or group category have been segregated into separate industries. Extensive product data are not collected for most "n. e. c." industries because of the small

---

[10] For the tires and inner tubes industry, data on materials consumption were used because they were considered better measures of output change than those based on quantity of product data, even though the quantity data yielded a coverage of more than 40 percent. See the notes on this industry in appendix C.

[11] This formula differs from that employed by Fabricant (Output, p. 370) in that the industry indexes indicated in the numerator and denominator of the right member are those with 1939 and 1947 weights, respectively, whereas under Fabricant's formula cross-weighted industry indexes are used in both numerator and denominator. The present formula is somewhat more precise (although the difference in most groups is practically insignificant) and Fabricant's represents an approximation made necessary by the fact that in his study only cross-weighted indexes were calculated for individual industries.

[12] See the 1947 census basic volumes, "General Explanations," paragraph headed "Transfer Values."

[13] In about one-third of the industries for which specific indexes were not calculated (measured in terms of 1947 value added), adequate data were collected for 1947 but comparable figures were not available for 1939. This is a rough indication of the increase in coverage that it probably will be possible to obtain in a future census index, comparing the year for which the next census is planned, 1953, with the year 1947. In the remaining cases, adequate data were not available for either 1947 or 1939.

importance of each of their products. A more important part of the missing area, however, consists of industries making highly diverse products, such as wiring devices and supplies, machine-shop products, food-products machinery, metal stampings, small leather goods, pharmaceutical preparations, prefabricated wood products, trimmings and art goods, jewelry, etc. The collection of meaningful quantity figures for most of these industries would require an amount of detail which was considered impractical.

TABLE 6.—NUMBER AND IMPORTANCE OF INDUSTRIES NOT REP-RESENTED BY SPECIFIC 1947 PRODUCTION INDEXES, BY IN-DUSTRY GROUP

		Missing industries				
		Num-ber of indus-tries	Value added (millions of dollars)		Percent of group value added	
No.	Title		1947	1939	1947	1939
	All manufacturing industries	190	18,301.8	5,254.1	24.6	21.5
20	Food and kindred products	7	888.2	230.6	9.8	6.6
21	Tobacco manufactures					
22	Textile mill products	8	546.0	203.5	10.2	11.4
23	Apparel and related products	18	984.6	312.9	22.2	22.1
24	Lumber and products, except furniture	11	814.4	281.6	32.6	35.7
25	Furniture and fixtures	15	1,231.0	365.6	89.4	87.5
26	Paper and allied products					
27	Printing and publishing industries					
28	Chemicals and allied products	13	1,460.1	476.7	27.0	26.2
29	Petroleum and coal products	2	31.2	13.7	1.5	2.4
30	Rubber products					
31	Leather and leather products	5	119.1	43.8	7.8	7.5
32	Stone, clay, and glass products	13	528.8	169.3	22.9	20.0
33	Primary metal industries	5	832.3	216.9	14.4	10.0
34	Fabricated metal products	18	2,337.3	660.9	47.5	47.3
35	Machinery (except electrical)	22	3,814.3	945.3	48.8	46.5
36	Electrical machinery	10	1,509.0	394.0	38.8	42.4
37	Transportation equipment	5	935.2	251.4	15.9	14.2
38	Instruments and related products	7	621.1	192.6	57.5	57.8
39	Miscellaneous manufactures	31	1,048.3	495.3	78.9	78.5

There are various possible ways of dealing with these missing industries. At one limit, they might simply be excluded from the calculations. The unadjusted group indexes could be specifi-cally described as relating only to the industries included, and similarly with a summary index, combining all unadjusted group indexes. At the other limit, estimates of the output change in each missing industry could be made on the basis of whatever information was available in the census volumes or elsewhere.

The first of these alternatives has the advantage of avoiding the need for estimates, which necessarily would be less reliable than the calculated indexes. But this advantage is largely fictitious. There are not likely to be many uses which would require composite indexes for the particular, and to some extent fortuitous, list of industries for which adequate data are available. If used at all, the composite indexes probably would most often be used as if they did in fact describe the output movements in entire groups and in all manufacturing. When this occurred, the user would be assuming implicitly that the industries included in the calculations were representative of all industries, or, to put it another way, that the average output change of missing industries was the same as that of the included industries. It was concluded, therefore, that if composite indexes were to be calcu-lated and used, it was best to consider the merits of this assump-tion relative to those of possible alternatives.

The second alternative—that of making individual estimates for each of the missing industries—would perhaps have been the most desirable to follow, if sufficient time and facilities had been available. In place of this detailed and laborious procedure, it was decided to make estimates, not for individual industries, but for all missing industries in each group taken together, and to use a uniform procedure in all groups.

Three alternatives were considered, all of which involved analogies between missing and included industries of individual groups. In general, this is a reasonable procedure because, under the classification system employed, industries which are similar in nature in one or more respects are grouped together. Many of the economic forces at work in individual industries would also affect other industries in the group to a greater or lesser extent, so that some degree of agreement might be expected in the changes that occur over time in particular variables.

One of the alternatives considered is mentioned above—to assume that the average output change calculated for included industries in each group was representative of missing industries as well. The other two involve the use of census statistics that are available for the missing industries—the data on value added and employment, respectively. The calculated change in value added or in employment for the sum of the missing industries of each group could form the basis for estimates of their output change, if information were at hand on changes for them in value added per unit of output, or in output per employee. (The change in value added, divided by the change in value added per unit of output, equals the change in output; and similarly, the change in employment, multiplied by the change in output per employee, equals the change in output.) The necessary infor-mation on value added per unit or on output per employee is not available for missing industries. However, such information is available, by calculation, for the included industries of each group, and the two alternative methods of estimating output changes for missing industries involve the assumptions that the calculated changes for included industries, taken together for each group, in value added per unit of output, or in output per employee, can be used as estimates of these changes in the group's missing industries.

Under all three methods, the estimated output indexes obtained for missing industries would be combined (with value added weights) with the "unadjusted" group indexes, relating to included industries, to obtain the "adjusted" group indexes. In the first of these methods, where the output changes are assumed to be similar, the adjusted and unadjusted group indexes would, of course, be identical. The second method, involving value added statistics, is mathematically equivalent to making a "coverage adjustment" to the unadjusted group indexes on the basis of the proportion of total value added in the group in each year accounted for by included industries. It is the method used by Fabricant.

The third method, involving employment data, is the one that was finally adopted. It is not equivalent to making a coverage adjustment on the basis of the proportions of total employment represented by included industries in the 2 years. The latter procedure amounts to deriving an index for missing industries by assuming a similar change in output per man, but then combining this with the index for included industries with employment rather than value added weights. Such a procedure would be proper only in cases where employment weights were desired, as, for example, in the type of productivity study referred to earlier in the section, "Definition of Physical Volume."

How different are the adjusted indexes obtained by these alternative methods? To determine this, all three were worked out, with results as shown in table 7. The spread in the all-manufacturing indexes is six points, or about 3 percent, with the output assumption yielding the lowest index, the output per man assumption the highest, and the value added assumption an intermediate figure. This ordering is due to the fact that the missing industries, in general, showed greater increases in employ-ment and in value added than did the included industries, with the margin somewhat greater for employment than for value added.

As might be expected, the range of difference varies in individual groups, depending on the importance of the missing sector (for four groups there are no missing industries, and no estimates are involved) and the relative changes in employment and value added for missing and included industries. In a number of cases the differences are small but in several, particularly chemicals and electrical machinery, they are rather large.

TABLE 7.—COMPARISON OF 1947 INDUSTRY-GROUP PRODUCTION INDEXES BASED ON ALTERNATIVE METHODS OF REPRESENTING "MISSING" INDUSTRIES (THOSE NOT REPRESENTED BY SPECIFIC INDEXES)

[1939=100]

No.	Title	Indexes obtained by assuming similarity for included and missing industries in average change in—			Percent of 1947 group value added accounted for by included industries
		Output	Value added per unit of output	Output per man	
	All manufacturing industries.....	178	181	184	75.4
20	Food and kindred products	150	155	159	90.2
21	Tobacco manufactures	155	155	155	100.0
22	Textile mill products	131	129	131	89.8
23	Apparel and related products	119	120	122	77.8
24	Lumber and products, except furniture	126	120	127	67.4
25	Furniture and fixtures	[1]	[1]	[1]	[1]
26	Paper and allied products	150	150	150	100.0
27	Printing and publishing industries	144	144	144	100.0
28	Chemicals and allied products	228	231	241	73.0
29	Petroleum and coal products	158	156	156	98.5
30	Rubber products	186	186	186	100.0
31	Leather and leather products	114	114	115	92.2
32	Stone, clay, and glass products	175	182	176	77.1
33	Primary metal industries	184	194	192	85.6
34	Fabricated metal products	200	200	201	52.5
35	Machinery (except electrical)	272	284	277	51.2
36	Electrical machinery	277	261	280	61.2
37	Transportation equipment	190	203	207	84.1
38	Instruments and related products	} [1] 177	[1] 185	[1] 193	[1] 23.0
39	Miscellaneous manufactures				

[1] Figures for group 25 are included with the figures for groups 38 and 39. For discussion of the special treatment of these three groups, see "Adjustments for missing industries."

NOTE.—Industry and group indexes employ 1939 weights.

Several tests were made in an effort to determine whether the changes in output, in value added per unit of output, or in output per man were most likely to be similar in a majority of cases for the included and missing industries in each group. As with the tests of the industry coverage adjustment described in appendix D, the calculations were necessarily confined to data for included industries, and the results therefore are not conclusive. The first investigation was of the degree of similarity—or rather, dissimilarity—in these variables among the included industries. The measure of dissimilarity, or variation, used was the standard deviation expressed as a percentage of the unweighted group average (the "coefficient of variation"). The results which are given in table 8 indicate that, in general, the output per man changes showed substantially less variation than either of the other two variables. Specifically, there was less variation in the output per man changes in every group but two, and in some cases, much less. The changes in output and value added per unit of output showed the least variation in only one group each.

These results were taken as rather strong evidence in support of the output per man assumption.

A second test was an empirical one, similar to a test used for the industry coverage adjustment that is described in appendix D. Briefly, groups were "redefined" to consist of the industries for which indexes were available, a number of indexes in each group were discarded, and new "experimental" indexes were calculated for the redefined groups using each of the three alternatives separately to estimate the output movements of the discarded industries. The experimental indexes were then compared with the original unadjusted group indexes, which included the discarded industries. Wherever possible the indexes discarded were for industries for which Fabricant had been unable to calculate indexes, on the theory that these "new additions" would most nearly resemble the indexes excluded in the present calculations.

The test was run twice, discarding a different set of indexes in each case, amounting to 28 and 25 percent, respectively, of the total value added for the redefined groups.[14] The results are shown in table 9, in which the headings "Test A" and "Test B" refer to the two test runs.

TABLE 8.—COEFFICIENTS OF VARIATION FOR 1947 INDEXES OF OUTPUT, VALUE ADDED PER UNIT, AND OUTPUT PER MAN, BY INDUSTRY GROUP

No.	Title	Coefficients of variation for industry indexes of—		
		Output	Value added per unit of output	Output per man
20	Food and kindred products	26.0	35.4	19.8
21	Tobacco manufactures	23.2	27.9	19.6
22	Textile mill products	25.2	23.3	14.3
23	Apparel and related products	29.6	14.8	10.4
24	Lumber and products, except furniture	34.7	23.6	5.7
25	Furniture and fixtures	[1]	[1]	[1]
26	Paper and allied products	5.2	19.4	3.8
28	Chemicals and allied products	51.7	55.5	20.0
29	Petroleum and coal products	46.1	30.4	22.8
30	Rubber products	9.0	24.7	3.8
31	Leather and leather products	15.0	20.2	16.7
32	Stone, clay, and glass products	30.2	20.2	24.2
33	Primary metal industries	19.5	15.1	8.8
34	Fabricated metal products	34.9	20.9	20.2
35	Machinery (except electrical)	35.3	27.0	15.1
36	Electrical machinery	40.4	21.2	9.4
37	Transportation equipment	38.8	17.6	7.7
38	Instruments and related products	} [1] 25.7	[1] 22.2	[1] 17.3
39	Miscellaneous manufactures			

[1] Figures for group 25 are included with the figures for groups 38 and 39.

NOTE.—Industry indexes employ 1939 weights. The deviations were computed from unweighted group means. Group 27—Printing and publishing industries—is not shown, since only a group index and no individual industry indexes were constructed.

[14] The furniture and printing groups were omitted from both test runs because only one industry is separately represented in the furniture group, and for printing only a group index and no individual industry indexes were constructed. Four additional groups were omitted from the second test run, because the small number of industries in the group left no alternatives to discard.

TABLE 9.—TEST OF ALTERNATIVE METHODS OF REPRESENTING "MISSING" INDUSTRIES IN CONSTRUCTING 1947 INDUSTRY-GROUP PRODUCTION INDEXES

[1939=100]

No.	Industry group Title	Original index for redefined group	Experimental indexes for redefined group based on assumed similarity between "discarded" and "retained" industries in—						Size of error (number of points difference between "experimental" and original indexes)					
			Output		Value added per unit of output		Output per man		Output		Value added per unit of output		Output per man	
			Test A	Test B	Test A	Test B	Test A	Test B	Test A	Test B	Test A	Test B	Test A	Test B
	All manufacturing industries:[1]													
	Test A	175	173	---------	176	---------	176	---------	-2	---------	+1	---------	+1	---------
	Test B	179	---------	180	---------	181	---------	180	---------	+1	---------	+2	---------	+1
20	Food and kindred products	150	145	149	142	155	147	150	-5	-1	-8	+5	-3	0
21	Tobacco manufactures	155	170	(1)	174	(1)	167	(1)	+15	(1)	+19	(1)	+12	(1)
22	Textile mill products	131	127	131	122	126	128	129	-4	0	-9	-5	-3	-2
23	Apparel and related products	119	124	118	121	119	121	120	+5	-1	+2	0	+2	+1
24	Lumber and products, except furniture	126	122	(1)	124	(1)	123	(1)	-4	(1)	-2	(1)	-3	(1)
26	Paper and allied products	150	148	(1)	157	(1)	151	(1)	-2	(1)	+7	(1)	+1	(1)
28	Chemicals and allied products	228	218	259	230	237	216	241	-10	+31	+11	+9	-12	+13
29	Petroleum and coal products	158	150	(1)	143	(1)	153	(1)	-8	(1)	-15	(1)	-5	(1)
30	Rubber products	186	177	189	193	192	182	187	-9	+3	+7	+6	-4	+1
31	Leather and leather products	114	112	114	126	114	113	116	-2	0	+12	0	-1	+2
32	Stone, clay, and glass products	175	173	186	182	182	182	170	-2	+11	+7	+7	+7	-5
33	Primary metal industries	184	181	180	194	194	191	187	-3	-4	+10	+10	+7	+3
34	Fabricated metal products	200	197	208	213	192	209	200	-3	+8	+13	-8	+9	0
35	Machinery (except electrical)	272	288	252	277	247	286	251	+16	-20	+5	-25	+14	-21
36	Electrical machinery	277	294	264	262	308	295	291	+17	-13	-15	+31	-11	+14
37	Transportation equipment	199	186	191	197	197	200	198	-13	-8	-2	-2	+1	-1
38, 39	Instruments and related products; Miscellaneous manufactures	196	196	208	205	201	191	211	0	+12	+9	+5	-5	+15

[1] Group 25—Furniture and fixtures and Group 27—Printing and publishing industries were omitted in both tests. In addition, Groups 21, 24, 26, and 29 were omitted in test B. The "all-manufacturing" indexes shown are with these omissions.

NOTE.—Indexes employ 1939 weights.

As can be seen from the table, in both test runs all three methods yielded relatively good estimates for the total, with the output per man method the best by a small margin. For the individual groups, the output per man method also led (in number of groups in each run for which the error was the smallest of the three) but involved errors of substantial size in several cases.

It was concluded from these investigations that the output per man assumption in most cases was likely to yield results closest to what would be found if output data were available for all industries, and this method was therefore used in developing the adjusted 1939–47 group indexes shown in table 1 and chapter IV. However, there is a measure of uncertainty attached to them, particularly in the groups where the alternatives showed important differences. Because the number of industries in the "included" and "missing" areas of each group is relatively small, averages of any variable for them are likely to be unstable, and in some cases would be dominated by the figures for one or two very important industries. This circumstance affects any general estimating method of the sort described, so that none can be relied on categorically to yield good, or even best, results in all cases.

An adaptation of the method described above was used for missing industries in the furniture and fixtures, instruments and related products, and miscellaneous manufactures groups. In these three groups, which together accounted for 6 percent of the value added in manufacturing in 1947, indexes were available for industries accounting for less than 40 percent of the value added in the group. For these groups the average output per man change in all other manufacturing groups was used to estimate output changes for missing industries, rather than the average change for included industries of the group. While this was done primarily because of the small number of industries in the three groups for which separate indexes were available, it would appear to be particularly appropriate for the miscellaneous group, the largest of the three. The alternative output estimates shown for these groups (combined) in table 7 are all based on relations (in output, value added per unit of output, and output per man) to the average of other groups, rather than to the calculated figures for included industries in these groups.

The employment figures used in the calculation of the adjusted group indexes were those for total employment, including force-account construction workers and administrative, sales, and office personnel, in addition to production workers. They are shown in appendix B. Unlike the data for production workers, total employment figures for 1939 had not been retabulated under the new industry classifications in the course of preparing the 1947 census basic volumes, and are shown here for all industries for the first time.[15]

It would have been preferable to use production workers only rather than total employment in this calculation, but this was not done because of some noncomparabilities in the breakdown of total employment between production workers and all others for 1939 and 1947. Differences in the language of the schedules used for the two censuses apparently led, in some 60 or 70 industries, to the reporting of certain types of workers as distribution employees in 1939 and as production workers in 1947; and to the reporting of other types of workers in some industries as force-account construction workers in 1939 and as production workers in 1947.

A further desirable refinement would have been to use man-hours worked by production workers, rather than number of workers, because of the possibility of differential changes in average hours worked between the included and missing industries of a group. However, the requisite data for 1939 are not available in the census for all industries.

[15] Because the cost of a special retabulation for the industries affected by classification changes was prohibitive, the total employment figures shown for these industries were partly estimated. The method was to distribute the employment figures for industries as originally defined for 1939 among industries corresponding to the definitions used for 1947 proportionately to the distribution of production workers found after retabulation, as shown in appendix D of the 1947 census basic volumes. Further adjustments were then made, where appropriate, corresponding to the revisions in the data for value of products and value added described earlier.